WTKSTES.IIi2K.17922.24

Machine Learning for space applications

Course project instruction

Instructions: Term papers to be attempted in groups of maximum two students per group. You can also make your solo group if you want. You need to submit your reports by 31st January. We will have short presentations in the first week of Feb. If it is OK with you all, we will record these and release them on the YouTube channel of STARS. Hence, we will prefer that the presentations are as sleek and cool as possible. Informative, accessible (like should not need a degree in EO) and less than 5 minutes. We can discuss it later.

Each term-paper will have its own set of requirements and objectives. In the end, we want a few things from all of you. The following rubric will help you to understand our core expectations.

Assessment criteria	**Description**	**Weight**
A. Scientific & Technical Rigor	Data preparation, model formulation, experiment design, validation methods.	30%
B. Implementation & Reproducibility	Functionality of code; organization of Colab notebook / GitHub repo; reproducible workflow.	20%
C. Analysis & Interpretation	Quality of results, clarity of evaluation metrics, critical discussion of model performance, awareness of limitations and data bias.	20%
D. Communication & Visualization	Clarity of figures, maps, legends, charts; appropriate color usage, readable legends, correct labeling. Logical narrative flow in report.	15%
E. Creativity & Insight	Originality of idea, thoughtful extension beyond minimum concept.	15%

Sentinel-5P Air Pollution Dynamics over Delhi (Last 24 Months)

Brief: Use Sentinel-5P TROPOMI data to build a 24-month time series of key gaseous pollutants over the Delhi NCR, focusing on tropospheric NO_2, SO_2, CO and formaldehyde (HCHO) as proxies for traffic/industrial emissions, power plants, biomass burning and secondary pollution chemistry. Generate monthly or weekly Level-3 composites over a

Delhi-centred region of interest, then couple these with reanalysis winds (e.g. ERA5) and simple trajectory or advection analysis to distinguish locally generated pollution from inflow transported from surrounding states and power-plant/biomass-burning regions. Apply hotspot and cluster analysis to identify persistent source regions (e.g. transport corridors, industrial belts, coal-based power clusters and seasonal crop-residue burning upwind of Delhi) and track the movement of high-pollution plumes into and out of the city under different synoptic conditions.

Expected outcomes:

- Gridded 24-month maps and animations of NO_2, SO_2, CO and HCHO over Delhi NCR, with seasonal anomaly maps highlighting winter smog periods and post-harvest burning episodes.nature+2
- Time-series plots of area-averaged column concentrations for each pollutant, decomposed by season and by “local” vs “advected” regimes based on wind direction and back-trajectory classification.acp.copernicus+1
- Source-region attribution maps showing typical upwind origin zones for severe episodes (e.g. northwest crop-burning areas, nearby industrial clusters, regional power plants) and typical transport pathways during prolonged smog events.eoportal+2
- Short interpretive note on Sentinel-5P’s strengths and limitations for Delhi air-quality studies, including column vs surface concentration issues and the impact of cloud screening and overpass time on source attribution.pmc.ncbi.nlm.nih+2

Learning outcomes:

- Practical experience in accessing and processing Sentinel-5P TROPOMI atmospheric products (NO_2, SO_2, CO, HCHO) for urban air-quality monitoring.dataspace.copernicus+2
- Understanding of how to combine satellite column data with meteorological fields to infer source regions and diagnose long-range transport vs local emissions in a megacity context.acp.copernicus+2
- Skills in spatio-temporal hotspot analysis and cluster-based characterization of pollution regimes, including interpretation of seasonal patterns and extreme events over Delhi.tandfonline+2
- Critical evaluation of the representativeness of satellite-derived column data for ground-level exposure, and of how open satellite and in-situ datasets can be jointly used for air-quality management and policy support in heavily polluted cities.pmc.ncbi.nlm.nih+2